

What is claimed is:

1. A method for executing a search in a peer-to-peer system, the method comprising:

receiving a query at a destination node;

5 receiving samples from a first set of nodes proximally located to the destination node in an overlay network for the peer-to-peer system, the samples associated with information stored at the proximally located nodes; and

identifying, based on the samples received from the first set of nodes, a first node of the first set of nodes likely storing information associated with objects
10 stored in the peer-to-peer system that are relevant to the query.

2. The method of claim 1, further comprising:

comparing the query to information stored in the first node; wherein the information stored in the first node is associated with objects stored in the peer-to-
15 peer network; and

generating search results including information stored in the first node associated with objects relevant to the query based on the comparison of the query to the information stored in the first node.

20 3. The method of claim 2, further comprising:

determining whether a quit threshold has been reached;

transmitting the search results to an initiator of the query in response to the quit threshold being reached; and

performing the following steps in response to the quit threshold not being reached:

identifying a second node likely storing information associated with objects stored in the peer-to-peer network that are relevant to the query based on samples received from a second set of nodes including the second node, wherein the second set of nodes are nodes proximally located to the first node in the overlay network; and

adding information stored in the second node to the search results; the added information being associated with objects that are relevant to the query.

4. The method of claim 3, wherein the quit threshold is associated with at least one of hops in the overlay network and whether the search results can be improved by adding information to the search results from the second node.

5. The method of claim 1, further comprising:

generating semantic vectors for objects stored in the peer-to-peer system;

hashing each of the semantic vectors to generate keys identifying locations in the overlay network for storing key-value pairs for the objects, wherein the keys are the semantic vectors for the objects and the values include at least one of the objects and addresses for the objects; and

storing the key-value pairs at nodes associated with the locations in the overlay network such that the stored key-value pairs associated with similar semantic vectors are proximally located in the overlay network.

5 6. The method of claim 5, further comprising:
 generating the samples for the first set of nodes as a function of at least one
 of key-value pairs stored at each of the first set of nodes.

 7. The method of claim 6, wherein generating the samples comprises:
10 generating a destination node semantic vector representative of objects
 associated with at least one of key-value pairs stored at the destination node and
 recent queries executed by the destination node;
 generating a list of key-value pairs for each node of the first set of nodes,
 wherein each list includes key-value pairs associated with objects having
15 semantics similar to the destination node semantic vector.

 8. The method of claim 7, wherein identifying, based on the samples received
 from the first set of nodes, a first node of the first set of nodes likely storing
 information associated with objects stored in the peer-to-peer network that are
20 relevant to the query comprises:
 generating a semantic vector for each of the samples for the first set of
 nodes;

comparing the destination node semantic vector to each of the semantic vectors for the first set of nodes; and

identifying one of the semantic vectors for the first set of nodes closest to the destination node semantic vector.

5

9. The method of claim 5, further comprising:

identifying lower elements for the semantic vectors;

generating planes in the overlay network associated with the lower elements; and

10 performing the steps of claim 1 for each of the planes.

10. The method of claim 5, further comprising:

storing indices of key-value pairs at the nodes;

15 replicating an index for a second node in the first node, wherein the second node is proximally located to the first node in the overlay network; and

identifying key-value pairs from the replicated index that are relevant to the query.

11. The method of claim 5, further comprising:

20 storing indices of key-value pairs at the nodes;

in the first node, replicating indices for a plurality of nodes in a region in the overlay network including the first node; and

identifying key-value pairs from the replicated indices that are relevant to the query.

12. The method of claim 2, wherein the first set of nodes are neighbor nodes to the destination node in the overlay network.

13. The method of claim 3, wherein the second set of nodes are neighbor nodes to the first node in the overlay network.

14. An apparatus for executing a search in a peer-to-peer system, the apparatus comprising:

means for receiving a query at a destination node;

means for receiving samples from a first set of nodes proximally located to the destination node in an overlay network for the peer-to-peer system, the samples associated with information stored at the proximally located nodes; and

means for identifying, based on the samples received from the first set of nodes, a first node of the first set of nodes likely storing information associated with objects stored in the peer-to-peer system that are relevant to the query.

15. The apparatus of claim 14, further comprising:

means for comparing the query to information stored in the first node;

wherein the information stored in the first node is associated with objects stored in the peer-to-peer network; and

means for generating search results including information stored in the first node associated with objects relevant to the query based on the comparison of the query to the information stored in the first node.

5 16. The apparatus of claim 15, further comprising:

 means for determining whether a quit threshold has been reached;

 means for transmitting the search results to an initiator of the query in response to the quit threshold being reached; and

 means for performing the following functions in response to the quit
10 threshold not being reached:

 identifying a second node likely storing information associated with objects stored in the peer-to-peer network that are relevant to the query based on samples received from a second set of nodes including the second node, wherein the second set of nodes are nodes proximally located to the
15 first node in the overlay network; and

 adding information stored in the second node to the search results; the added information being associated with objects stored in the peer-to-peer system that are relevant to the query.

20 17. The apparatus of claim 16, wherein the quit threshold is associated with at least one of hops in the overlay network and whether the search results can be improved by adding information to the search results from the second node.

18. A computer readable medium on which is embedded a program, the program performing a method, the method comprising:

receiving a query at a destination node;

receiving samples from a first set of nodes proximally located to the destination node in an overlay network for the peer-to-peer system, the samples associated with information stored at the proximally located nodes; and

identifying, based on the samples received from the first set of nodes, a first node of the first set of nodes likely storing information associated with objects stored in the peer-to-peer system that are relevant to the query.

19. The computer readable medium of claim 18, wherein the method further comprises:

comparing the query to information stored in the first node; wherein the information stored in the first node is associated with objects stored in the peer-to-peer network; and

generating search results including information stored in the first node associated with objects relevant to the query based on the comparison of the query to the information stored in the first node.

20. The computer readable medium of claim 19, wherein the method further comprises:

determining whether a quit threshold has been reached;


transmitting the search results to an initiator of the query in response to the quit threshold being reached; and

performing the following steps in response to the quit threshold not being reached:

5 identifying a second node likely storing information associated with objects stored in the peer-to-peer network that are relevant to the query based on samples received from a second set of nodes including the second node, wherein the second set of nodes are nodes proximally located to the first node in the overlay network; and

10 adding information stored in the second node to the search results; the added information being associated with objects stored in the peer-to-peer system that are relevant to the query.

21. The computer readable medium of claim 20, wherein the quit threshold is
15 associated with at least one of hops in the overlay network and whether the search results can be improved by adding information to the search results from the second node.

22. A peer-to-peer system comprising: 
20 a plurality of nodes in the system operating as a search engine operable to execute a query received by the search engine;
an overlay network implemented by the plurality of nodes;

a plurality of indices stored at the plurality of nodes, each index including at least one semantic vector for an object;

wherein a first node in the search engine is operable to receive samples from nodes proximally located to the first node in the overlay network, the first node utilizing the samples to identify an index of one of the other nodes to search in response to receiving the query.

23. The system of claim 22, wherein similar semantic vectors are stored at nodes proximally located in the overlay network.

24. The system according to claim 23, wherein the first node is located in a region in the overlay network and the first node is operable to store indices from nodes in the region, such that the first node is operable to search a plurality of indices likely to include information relevant to the query without forwarding the query to other nodes in the region.